Cyber Warfare: Steganography vs. Steganalysis

BY HUAIQING WANG AND SHUOZHONG WANG

For every clever method and tool being developed to hide information in multimedia data, an equal number of clever methods and tools are being developed to detect and reveal its secrets.

The rise of the Internet and multimedia techniques in the mid-1990s has prompted increasing interest in hiding data in digital media. Early research concentrated on watermarking to protect copyrighted multimedia products (such as images, audio, video, and text) [1, 8]. Data embedding has also been found to be useful in covert communication, or steganography. The goal was and still is to convey messages under cover, concealing the very existence of information exchange.

Compared to watermarking, steganography has drawn less attention until recently, as computer specialists, signal-processing researchers, and multimedia product vendors concerned about information security have recognized that illicit use of the technique might become a threat to the security of the worldwide information infrastructure [6]. Researchers have thus begun to study steganalysis, or the detection of embedded information. Detecting secret data hidden in millions of multimedia items downloadable from online sites is recognized as an especially difficult task [10].

The idea and practice of hiding information exchange has a long history. Traditional techniques of steganography, or covered writing in Greek, ranged from tat-

tooing the shaved head of a trusted messenger during ancient times (as reported by the 5th century Greek historian Herodotus) to using invisible ink during the two World Wars in the 20th century. Modern steganography employs digital media content as camouflage, powerful computers and signal-processing techniques to hide secret data, and methods to distribute stego-media throughout cyberspace, thus posing a serious challenge to scientists and professionals alike in the field of information security.

The two major branches of information hiding, steganography and watermarking, share many characteristics. They also differ in a number of ways, including purpose, specifications, and detection/extraction methods (see Table 1). The most fundamental difference is that the object of communication in watermarking is the host signal, with the embedded data providing copyright protection. In steganography the object to be transmitted is the embedded message, and the cover signal serves as an innocuous disguise chosen fairly arbitrarily by the user based on its technical suitability. In addition, the existence of a watermark is often declared openly, and any attempt to remove or invalidate the embedded content renders

ILLUSTRATION BY ISTVÁN OROTZ

the host useless. The crucial requirement for steganography is perceptual and algorithmic undetectability. Robustness against malicious attack and signal processing is not the primary concern, as it is for watermarking.

Steganography also differs from cryptography, which does not conceal the communication itself but only scrambles the data to prevent eavesdroppers understanding the content. Cryptography involves various methods and implementations. Steganography, on the other hand, is a relatively new area of study, as reflected in the research focus of published papers. Steganography and cryptography may be considered complementary and orthogonal. Anyone engaging in secret communication can always apply a cryptographic algorithm to the data before embedding it to achieve additional security. In any case, once the presence of hidden information is revealed or even suspected, the purpose of steganography is defeated, even if the message content is not extracted or deciphered.

	Requirements	Watermarking		Steganography
		Private	Public	
Goal	Protection of intellectual property rights	++++		–
	Transmission of secret message without raising suspicion	–		++++
Specifications	Perceptual invisibility	++++		+++++
	Statistical or algorithmic invisibility	+		+++++
	Robustness against hostile removal, destruction, or counterfeiting	+++++		–
	Resistance against normal signal processing	++++		+
	Capable of surviving common compression coding	++++		++
	Large payload	++		++++
Detection/ Extraction	Extractability/detectability without host/cover object	–	++++	++++
	Extractability only with presence of host/cover object	++++	–	–
	Requirement of low complexity in extraction/detection	++		+++
	Optional capability of automatic object downloading	+		++

Note: Crucial: +++++ Necessary: ++++ Important: +++ Desirable: ++ Useful: + Unnecessary or irrelevant:: –
Public watermarking schemes do not need the host signal in detection/extraction; private schemes require the presence of the host.

Table 1. Steganography and watermarking.

Steganographic Techniques

Images are the most popular cover media for steganography and can be stored in a straightforward bitmap format (such as BMP) or in a compressed format (such as JPEG). Palette images are usually in the GIF format. Information hiding is accomplished either in the space domain or in the frequency domain. In terms of insertion schemes, several methods (such as substitution, addition, and adjustment) can be used. One adjustment approach is Quantization Index Modulation (QIM), which uses different quantizers to carry different bits of the secret data [2]. Although a simple unified method for classifying these techniques does not exist, some popular approaches are used in downloadable steganographic tools or found in the literature (see Table 2).

EzStego	online.securityfocus.com/tools/586/scoreit/
F5	wwwrn.inf.tu-dresden.de/~westfeld/f5.html
Hide and Seek v4.1	ftp://ftp.csua.berkeley.edu/pub/cypherpunks/steganography/
Hide and Seek for Win95	ftp.hacktic.nl/pub/crypto/incoming/
Hide4PGP	www.heinz-repp.onlinehome.de/Hide4PGP.htm
Jpeg-Jsteg	ftp://ftp.funet.fi/pub/crypt/steganography/
Mandelsteg	ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/
MP3Stego	www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/
OutGuess	www.outguess.org/download.php
Steganos	www.steganos.com/en/
S-Tools v4	members.tripod.com/steganography/stego/s-tools4.html
White Noice Storm	ftp://ftp.esua.berkeley.edu/pub/cypherpunks/steganography/

For a steganography tool table, see
www.jjtc.com/Steganography/toolmatrix.htm

Table 2. Steganography tools and their sources.

LSB modification. These techniques are based on modifying the least significant bits (LSBs), of the pixel values in the space domain. In a basic implementation, these pixels replace the entire LSB-plane with the stego-data; on average, 50% of the LSBs are flipped (see Figure 1). It can be shown that fidelity of the stego-image measured in peak-signal-to-noise ratio with respect to the cover is 51.1dB, representing a very high degree of imperceptibility compared to the lower bound of 39dB generally accepted by researchers of watermarking. With more sophisticated schemes in which embedding locations are adaptively selected, depending on human vision characteristics, even less distortion is achievable. Popular tools include EzStego, S-Tools, and Hide and Seek. In general, simple LSB embedding is susceptible to image processing, especially lossy compression.

Masking approaches. These techniques are similar to visible watermarking in which pixel values in masked areas are raised or lowered by some percentage. Reducing the increment to a certain degree makes the mark invisible. In the patchwork method, pairs of patches are selected pseudo-randomly; pixel values in each pair are raised by a small constant value in one patch and lowered by the same amount in the other.

Transform domain techniques. Data embedding performed in the transform domain is widely used for robust watermarking. Similar techniques can also realize large-capacity embedding for steganography. Candidate transforms include discrete cosine transform (DCT), discrete wavelet transform (DWT), and dis-

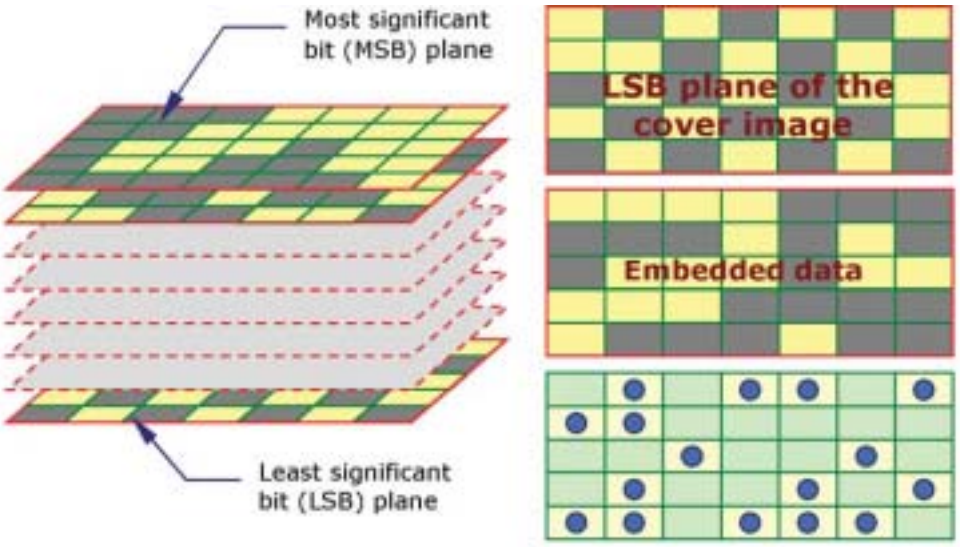

Figure 1. A basic LSB approach. Bit-planes of a grayscale image are sketched on the left with MSB on top. Dark and light boxes represent binary values 0s and 1s, respectively, of the pixels on different bit-planes. The LSB-plane of the cover image on the top right is replaced with the hidden data in the middle, which becomes the LSB-plane of the stego-image. The bottom-right map indicates differences between LSB planes of the cover- and stego-images. Circles represent the flipped bits; with an average of 50% bits in the LSB plane changed, the stego-image is visually identical to the cover.

crete Fourier transform (DFT). By being embedded in the transform domain, the hidden data resides in more robust areas, spread across the entire image, and provides better resistance against signal processing. Various methods are available. For example, we can perform a block DCT and, depending on payload and robustness requirements, choose one or more components in each block to form a new data group that, in turn, is pseudo-randomly scrambled and undergoes a second-layer transformation. Modification is then carried out on the double transform domain coefficients using various schemes.

Techniques incorporated in compression algorithms. The idea is to integrate the data-embedding with an image-compression algorithm (such as JPEG). For example, the steganographic tool Jpeg-Jsteg takes a lossless cover-image and the message to be hidden to generate a JPEG stego-image. In the coding process, DCT coefficients are rounded up or down according to individual bits to be embedded. Such techniques are attractive because JPEG images are popular on the Internet. Other transforms (such as DFT and wavelet transform) can also be used.

Spread-spectrum techniques. The hidden data is spread throughout the cover-image based on spread-spectrum techniques (such as frequency hopping). A stego-key is used for encryption to randomly select the frequency channels. White Noise Storm is a popular tool using this technique. In other research [7], with embedded data as the object to be transmitted, the cover-image is viewed as interference in a covert communication framework. The embedded data is first modulated with pseudo-noise so the energy is spread over a wide frequency band, achieving only a very low level of embedding strength. This is valuable in achieving imperceptibility.

The three most important requirements that must be satisfied for any steganographic system are: security of the hidden communication; size of the payload; and robustness against malicious and unintentional attacks.

Security. In order to avoid raising the suspicions of eavesdroppers, while evading the meticulous screening of algorithmic detection, the hidden contents must be invisible both perceptually and statistically [5]. Some information-theoretic-based definitions for a perfectly secure system assume detailed knowledge of the statistics of the cover and require unlimited computational resources. These conditions cannot be strictly met in real-world steganographic applications. For example, regarding statistical knowledge, one may be able to estimate the statistics of a particular ensemble of signals frequently used by a certain group of people and establish a model for detection. But such models are meaningless if the estimation error exceeds the extent of modifications caused by embedding. Moreover, the computational complexity of any useful steganalytic tools cannot be infinitely great. In terms of practicality, a system may be considered secure, or steganographically strong [9], if it is impossible for an eavesdropper to detect the presence of stego-contents by using any accessible means.

Payload. Unlike watermarking, which needs to embed only a small amount of copyright information, steganography aims at hidden communication and therefore usually requires sufficient embedding capacity. Requirements for substantial data capacity and security are often contradictory. Depending on specific application scenarios, a trade-off must be sought.

Robustness. Although robustness against attacks is not a top priority, as in watermarking, being able to withstand JPEG coding is certainly desirable, since most true-color images are JPEG-compressed before being put online.

a b c d

Figure 2. Effects of simple LSB embedding: (a) the stego-image is visually indistinguishable from the cover; (b) the LSB-plane of the cover-image in which some features are visible corresponds to areas with smooth and nearly saturated colors; (c) the LSB-plane of the stego-image contains embedded data (0.52 bits per color channel) scrambled and evenly spread, so the original features are blurred; and (d) the LSB-plane of the stego-image results from sequential embedding, and only the red and part of the green components are filled with data, while the others are padded with os.

Detection of Steganographic Content

Despite the fact that steganographic tools alter only the least-significant image components, they inevitably leave detectable traces in the stego-image, so successful attacks are still possible [4]. The primary goal of attacks against steganography is the detection of the presence of hidden data, although in some cases it may also include extraction and/or destruction of the data. Here, steganalysis refers to detection of the presence of hidden information in a given image. Assume, too, that the cover-image is not available to the steganalyst (stego-only detection). In general, steganalysis involves two major types of techniques: visual analysis and statistical (algorithmic) analysis.

Steganalytic Methods	Description	Targeted Steganographic Techniques
RS steganalysis	Sensitivity of dual statistics based on spatial correlation of pixels to LSB randomization due to steganographic embedding is used in analysis.	Various LSB modification techniques
PoV-based Chi-square test	A Chi-square test checks whether the occurrence of each pair of values tends to become equal, indicating some data is embedded.	Steganography based on swapping pairs of values of pixel gray levels, colors, or DCT coefficients
Palette checking	Peculiarity in palette ordering is a clear sign of systematic modification.	Steganography in palette images
RQP method	Method based on analyzing the increased number of close-color pairs caused by embedding.	LSB embedding in true-color images
Check JPEG compatibility	Method detects unusual departure from the JPEG signature inherent in images initially stored in JPEG format.	Space-domain steganography using images initially stored in the JPEG format
Histogram analysis	Method reveals discreteness or periodicity in particular coefficients due to quantization-related modification.	QIM or other quantization-related embedding methods
Universal blind detection	Statistical quantities constructed using high-order statistics, and a detection model established with the threshold obtained in a training process.	Various steganographic techniques

Table 3. Popular steganalytic methods.

Visual analysis tries to reveal the presence of secret communication through inspection, either with the naked eye or with the assistance of a computer. The computer can, for example, help decompose an image into bit-planes. Any unusual appearance in the display of the LSB-plane would be expected to indicate the existence of secret information. Visual inspection can succeed when secret data is inserted in relatively smooth areas with pixel values near saturation.

Statistical analysis is more powerful since it reveals tiny alterations in an image's statistical behavior caused by steganographic embedding. As there is a range of approaches to embedding, each modifying the image in a different way, unified techniques for detecting hidden information in all types of stego-images are difficult to find. The nominally universal methods developed to detect embedded stego-data are generally less effective than the steganalytic methods aimed at specific types of embedding.

Figure 2 illustrates the effects of a simple LSB steganographic operation. The stego-image in (a) is visually identical to the cover-image. The LSB-plane of the cover-image in (b) contains some noticeable features corresponding to the areas with flat and saturated colors; it is partially replaced with embedded data, scrambled, and spread over the bit-plane, as in (c), with an embedding capacity of 0.52 bits per color channel. The original features are blurred, as they are masked by the embedded data. If the data is embedded sequentially, the result is the map in (d), whereby the message data occupies only the red and a portion of the green planes, with the rest of the LSB-plane padded with zeros. More sophisticated embedding

schemes (such as selecting busy areas and padding unoccupied space with a random sequence with the same statistical property as the cover-image) make direct analysis of the LSB-plane more difficult.

For palette-based images (such as those in GIF format), replacing LSB with the embedded data causes significant color singularities, since neighboring indices in a palette may point to different colors. In order to avoid this problem, some stego tools (such as EzStego) rearrange the palette so consecutive indices represent similar colors. Other techniques simply shuffle the palette according to a key, with the image itself staying intact. These methods can be defeated through analysis of the palette to find unusual ordering, as normal commercial software products arrange the palette based on such attributes as color components and luminance. A peculiar palette itself is sufficient to arouse a steganalyst's suspicion.

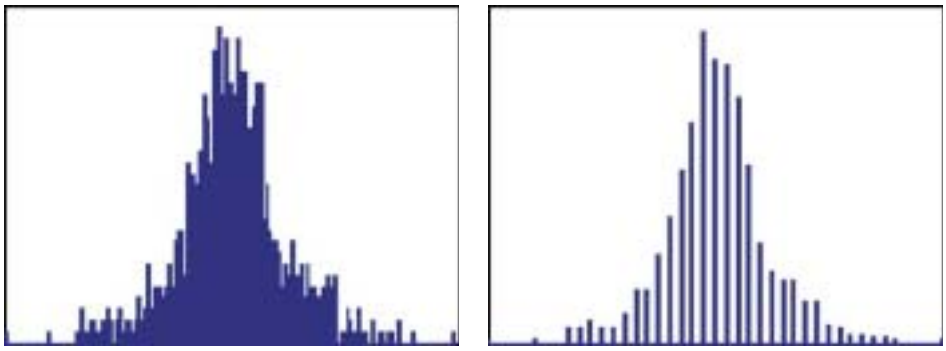

Figure 3. Histogram analysis: (left) histogram of dual-transform coefficients of a cover-image Lena; (right) the signature of QIM embedding.

With a steganographic tool at hand, the choice of cover-images is at the user's discretion. Images stored in the JPEG format (abundant due to the bandwidth limitation and the widespread use of digital cameras) are therefore likely to be chosen as covers for hidden communication. However, because the JPEG algorithm performs quantization on block DCT coefficients, some known structures are inherent in these images. Slight modification therefore leaves traces incompatible with the JPEG signature, making the stego-image vulnerable to analysis [3].

Some embedding tools cause subtle changes in the set of possible values that may be taken by the pixel gray levels and/or transform coefficients. For example, the QIM technique uses a number of quantizers to embed data into the cover-image so the sample values or certain coefficients show signs of discreteness. A histogram analysis may be used to reveal such a signature (see Figure 3); the left histogram represents the distribution of a particular group of coefficients taken from a cover-image generated by the double transform scheme described earlier. In a stego-image, the discreteness of the histogram on the right is a clear sign of QIM embedding.

Because detecting stego content is performed only with current steganalytic approaches, any system considered secure today may be broken tomorrow using new techniques. Some of the most popular steganalytic methods are outlined in Table 3.

Blind detection of hidden information in apparently innocuous digital media is generally more challenging than data embedding, especially when the embedding rate is low [12], as steganalysts always work in passive mode. Another important consideration in steganalysis is to keep the computational complexity sufficiently low, allowing the screening of thousands (even millions) of suspected images in a reasonably short amount of time. The computation limitation may be less stringent for steganography, since in practical applications the embedding algorithm is executed on only a few images taken from a large database.

Conclusion

The battle between steganography and steganalysis represents an important part of 21st century cyber warfare with a profound influence on information security. The two sides of the battle are the attempt to transmit secret messages under cover of innocuous multimedia signals and the effort to detect or prevent such hidden communication.

Various steganographic tools have been developed, many available online. In a sense, some simple methods are already defeated due to the relentless endeavor of steganalysts. Meanwhile, countermeasures against steganalysis are also emerging [11]. Tools that can withstand, to some degree, both visual and statistical attacks are being introduced. For example, in data embedding, much effort has gone toward preserving the statistical characteristics of the cover media. To combat steganalytic tools based on analyzing the increase of unique colors in an image, new embedding methods may be devised that avoid creation of new colors. Alternatively, modifications leading to detectable artifacts may be compensated for after embedding while ensuring the intended recipient is still able to extract the secret message.

Apart from their law enforcement/intelligence and

anti-terrorist significance, steganographic techniques also have peaceful applications, including: in-band captioning; integration of multiple media for convenient and reliable storage, management, and transmission; embedding executables for function control; error correction; and version upgrading. Computer specialists, signal-processing researchers, and information security professionals should expect to devote much more attention to the challenging area of information hiding and detection. C

References

1. Berghel, H. and O'Gorman, L. Protecting ownership rights through digital watermarks. *IEEE Comput. 29,* 7 (July 1996), 101–103.
2. Chen, B. and Wornell, G. Quantization index modulation: A class of provably good methods for watermarking and information embedding. *IEEE Transact. Info. Theory 47,* 4 (May 2001), 1423–1443.
3. Fridrich, J. and Goljan, M. Practical steganalysis of digital images: State of the art. In *Proceedings of SPIE, Security and Watermarking Multimedia Contents IV* (San Jose, CA, Jan. 21–24). International Society for Optical Engineering, 2002, 1–13.
4. Johnson, N. and Lajodia, S. Exploring steganography: Seeing the unseen. *IEEE Comput. 31,* 2 (Feb. 1998), 26–34.
5. Katzenbeisser, S. and Petitcolas, F. Defining security in steganographic systems. In *Proceedings of SPIE, Security and Watermarking of Multimedia Contents IV* (San Jose, CA, Jan 21–24). International Society for Optical Engineering, 2002, 50–56.
6. Kovacich, G. and Jones, A. What infosec professionals should know about information warfare tactics by terrorists (Parts 1 and 2). *Comput. & Sec. 21,* 1 (Jan. 2002), 35–41; *21,* 2 (Mar. 2002), 113–119.
7. Marvel, L., Boncelet, C., Jr., and Retter, C. Spread-spectrum image steganography. *IEEE Transact. Image Process. 8,* 8 (Aug. 1999), 1075–1083.
8. Memon, N. and Wong, P. Protecting digital media content. *Commun. ACM 41,* 7 (July 1998), 34–43.
9. Moskowitz, I., Longdon, G., and Chang, L. A new paradigm hidden in steganography. In *Proceedings of the 2000 Workshop on New Security Paradigms* (Ballycotton, County Cork, Ireland, Sept. 18–21). ACM Press, New York, 2000, 41–50.
10. Provos, N. and Honeyman, P. Detecting steganographic content on the Internet. In *Proceedings of Network and Distributed System Security Symposium* (San Diego, Feb. 6–8). Internet Society, Reston, VA, 2002.
11. Westfeld, A. F5-Steganographic algorithm: High capacity despite better steganalysis. In *Lecture Notes in Computer Science 2137.* Springer-Verlag, Berlin, 2001, 289–302.
12. Westfeld, A. Detecting low embedding rates. In *Lecture Notes in Computer Science 2137.* Springer-Verlag, Berlin, 2003, 324–339.

Huaiqing Wang (iswang@is.cityu.edu.hk) is an associate professor in the Department of Information Systems at the City University of Hong Kong.

Shuozhong Wang (shuowang@staff.shu.edu.cn) is a professor in the School of Communication and Information Engineering at Shanghai University, China.

This research is supported by research grant No. CityU 1234/03E from the government of Hong Kong.